AI Market Research
Sentiment Analysis at Google Nest and Amazon Echo
Track B: Regulator

Prashansa Atri
Leeds School of Business
University of Colorado,
Boulder
Prashansa.Atri@colorado.edu

Elise Thomason Larsen
Leeds School of Business
University of Colorado,
Boulder
Elise.ThomasonLarsen@colorado.edu

Jesse Ihns
Leeds School of Business
University of Colorado,
Boulder
Jesse.Ihns@colorado.edu

Geo Vadakkan
Leeds School of Business
University of Colorado,
Boulder
Geo.Vadakkan@colorado.edu

Abstract

Sentiment analysis is a computational approach that aims to extract and analyze subjective information from textual data. With the exponential growth of online communication and social media platforms, sentiment analysis has become increasingly relevant in understanding people's opinions, attitudes, and emotions towards various topics, products, or services. It utilizes natural language processing (NLP) techniques and machine learning algorithms to process and analyze large volumes of data (text, expressions, spoken words etc.).

By examining the sentiment expressed in customer reviews, social media posts, surveys, or other textual sources, sentiment analysis provides valuable insights for businesses, marketers, and decision-makers. It enables them to understand public perception, customer satisfaction, brand reputation, and market trends.

However, sentiment analysis has various implications that need to be carefully considered. Ethical concerns related to privacy, data protection, and the responsible use of personal information arise. Fairness and avoiding bias in sentiment analysis models are critical to prevent reinforcing stereotypes or discriminatory practices. Proper regulation and transparency are necessary to address these concerns.

1. Introduction

Over the past decade and a half, an increasing number of statements have been made online — a result of, among other things, the adoption of both social media and eCommerce platforms. Progress surrounding the deployment of CNNs, RNNs and LSTMs for unsupervised NLP tasks has greatly improved opinion/sentiment extraction [13]. As a result, large tech firms that provide consumer facing products and services are now in a position to monetize these sentiment extractions. Specifically, they can create and track cross-sections that connect the following:

1. the **subject** (consumer with a unique ID), who
2. expresses a **sentiment** about
3. an **object** (product, service, politician etc.).

From there, the large tech firms tracking these instances of consumer sentiment can follow up with targeted actionables; they can, given a psychographic profile partially formed by these expressed sentiments, recommend specific products, services or promotional content. Accordingly, concerns surrounding consumer welfare now make themselves apparent — are consumer privacy and data security important considerations that are informing the decisions made?

As prospective regulators considering the present – and possible future – extension of sentiment analysis, it is important we understand the current environment. Therefore, the technical fundamentals that make sentiment analysis possible will be explored. In the process, the ways this technology has been and is deployed by Google and Amazon will be discussed. These findings were then incorporated with a market survey that recorded and analyzed consumer preferences as it concerned the privacy/convenience tradeoff that sentiment analysis may represent. With the help of a data scientist working at a large tech firm on AR/VR, these findings will be contextualized. Next, there will be a review of existing literature that grapples with the ethical tensions made apparent by other, tangentially related, deployments of deep learning methods. Finally, as regulators, the conclusion will contain specific recommendations that best balance the central tension between innovation and consumer welfare.

2. Research on the Area of Application

Sentiment analysis has gained significant traction globally in recent years, with the market size expected to reach **USD 10.1 billion** by 2030, growing at a CAGR of **14.2%** from 2022 to 2030, according to a report by ReportLinker [1].

The rise of social media, e-commerce, and online customer feedback platforms has contributed to the growth of sentiment analysis, as businesses seek to understand and analyze customer opinions and attitudes towards their products, services, and brands.

Sentiment analysis has become an essential tool for businesses in various industries, including retail, healthcare, hospitality, and finance, among others, as it helps them to identify trends, patterns, and insights that can inform their decision-making processes.

In addition, the development of advanced machine learning and natural language processing technologies has enabled sentiment analysis to become more accurate and efficient, with the ability to analyze large volumes of data in real-time.

However, there are still some challenges facing sentiment analysis, such as the difficulty of accurately interpreting the nuances of human emotions and the potential biases that can arise from using automated algorithms. Nevertheless, sentiment analysis is expected to continue to grow in importance and adoption in the coming years, as businesses increasingly seek to gain a competitive edge through understanding and responding to customer sentiment.

There are several reasons why businesses and researchers may want to conduct research on sentiment analysis:

Improve customer experience: By understanding customer sentiment, businesses can identify areas where they need to improve their products or services to better meet customer needs and expectations. This can lead to higher customer satisfaction and loyalty.

Identify brand reputation issues: Sentiment analysis can help businesses identify negative sentiment towards their brand or products and take appropriate action to address those issues, thereby protecting their reputation and improving customer trust.

Inform marketing strategies: Sentiment analysis can provide insights into customer preferences and opinions, allowing businesses to tailor their marketing messages and campaigns to better resonate with their target audience.

Competitive analysis: Sentiment analysis can also be used to compare the sentiment towards a business or brand with that of its competitors, providing valuable insights into the relative strengths and weaknesses of different brands in the marketplace.

Risk management: Sentiment analysis can be used to monitor and analyze public sentiment towards a company or industry, allowing businesses to anticipate and mitigate potential risks or crises.

Product development: By analyzing customer feedback and sentiment towards existing products, businesses can identify areas for product improvement or new product development opportunities.

Sentiment analysis is offered as a service to other companies or used in-house to gain insights into customer sentiment and behavior. Some potential revenue streams for a sentiment analysis business model include:

Subscription model: Companies can offer sentiment analysis tools and dashboards as a subscription-based service to customers. They can charge a monthly or yearly fee based on the number of data points analyzed or the level of features and insights provided.

Pay-per-use model: Another option is to charge customers based on the number of sentiment analysis requests they make. This model is particularly useful for smaller businesses or those with irregular data analysis needs.

Consulting services: Companies can offer consulting services to help clients interpret and act on the insights generated by sentiment analysis. This can include developing marketing strategies, improving customer service, or optimizing product offerings.

API integration: Companies can offer sentiment analysis as an API service to other businesses, allowing them to integrate sentiment analysis into their own software applications.

Data analysis: Sentiment analysis companies can collect and analyze data on customer sentiment, behavior, and trends and sell this information to third-party organizations for market research purposes.

Sentiment analysis typically involves the use of various AI technologies to analyze text data and determine the sentiment expressed within it. Some of the most used AI technologies for sentiment analysis include:

Natural Language Processing (NLP): NLP is a branch of AI that deals with the interaction between computers and human language. It helps sentiment analysis tools to identify the meaning of words and phrases in a sentence, which is essential for determining the overall sentiment expressed.

Machine Learning: Machine learning algorithms are used to train sentiment analysis models to identify sentiment in text data. By analyzing large volumes of text data, these models can learn to recognize patterns and relationships that indicate positive, negative, or neutral sentiment.

Deep Learning: Deep learning is a subset of machine learning that uses artificial neural networks to analyze data. Deep learning algorithms are particularly effective at processing complex data structures such as text, making them well-suited for sentiment analysis tasks.

Sentiment Lexicons: Sentiment lexicons are pre-built dictionaries of words and phrases that are associated with positive, negative, or neutral sentiment. Sentiment analysis tools can use these lexicons to identify the sentiment of words and phrases within text data.

Emotion Detection: Emotion detection algorithms use NLP and machine learning techniques to identify emotional states expressed within text data. Emotion detection can be particularly useful in analyzing customer feedback or social media posts to determine the emotional state of the customer.

Topic Modeling: Topic modeling is a technique used to identify the main topics discussed within a piece of text data. By understanding the context of the text, sentiment analysis tools can more accurately determine the sentiment expressed towards a particular topic.

Overall, the sentiment analysis business model can be profitable, provided that the company can offer a high level of accuracy and insight while keeping costs low. Additionally, building strong partnerships with other businesses and developing innovative features and services can help companies stand out in an increasingly competitive market.

What was most intriguing is that sentiment analysis can help businesses gain a better understanding of their customers and the market and make data-driven decisions that can lead to improved business outcomes. So, to learn more about sentiment analysis, the research scope was narrowed to research on sentiment analysis in Google Analytics and Amazon Echo, which involves the study of customer opinions, emotions, and attitudes towards products, services, or brands.

In an effort to identify the customer behavior related to sentiment analysis, a market survey was conducted on **124** people, ages **15** and older. Out of these respondents, **66%** respondents were from India, **27%** were from the United States and the final **7%** were from the U.K.,

Australia, and Canada. The respondents were made up of **38%** females and **62%** males **(Appendix 1)**. From the analysis of this survey, it was evident that although **78%** of people support technological advancements, only **38%** are ready to sacrifice their privacy to make themselves more self-sufficient with technology **(Appendix 2)**.

The survey also revealed that **49%** of the respondents believe sentiment analysis warrants additional data when compared to speech and text. An additional **52%** are ready to allow the use of sentiment analysis to promote content which is valuable for their emotional status at any moment **(Appendix 3)**.

Respondents' views on the government's stance on sentiment analysis reveals that **67%** believe that the government has the capability to regulate the technology, but only **30%** feel that they are taking necessary steps to make it more secure for the end user **(Appendix 4)**.

3. Company Comparisons – Solutions and Competitive Positioning

Google Nest is a line of smart home devices developed by Google. The Nest product line includes a variety of smart home devices such as smart thermostats, cameras, video doorbells, smart speakers, and more.

Nest products are designed to work together seamlessly, allowing users to control all their smart home devices through a single interface. For example, users can control their Nest thermostat using the Nest app or with their voice through a Google Nest Mini smart speaker.

The Nest devices also use machine learning to adapt to users' behaviors and preferences over time, allowing them to save energy and be more efficient. For example, the Nest Learning Thermostat can learn users' preferred temperature settings and adjust them automatically based on the time of day and whether anyone is home.

Amazon Echo is a smart speaker device that uses voice recognition and artificial intelligence to interact with users, providing information, entertainment, and home automation features. Sentiment analysis is applied to Amazon Echo data to analyze the tone and emotions of user interactions, based on their voice commands, questions, and feedback.

Market research on sentiment analysis in Google Analytics and Amazon Echo helps businesses understand customer needs, preferences, and pain points, as well as improve their products, services, and marketing strategies.

It can also help businesses identify and address negative feedback, complaints, or issues that may affect their reputation and customer loyalty.

A review of the market survey revealed that **48%** of respondents trust Google products with their data more, which compares to **23%** of people who trust Amazon more. **27%** of respondents trusted neither **(Appendix 5).**

Latest Technologies Used by Amazon ECHO and Google NEST in context of Sentiment Analysis.

Amazon Echo uses the following technologies:

Alexa Voice Service (AVS) : AVS uses Automatic Speech Recognition (ASR) to convert spoken words into text, and Natural Language Understanding (NLU) to interpret the meaning behind the text [2]. It leverages deep learning models and other machine learning algorithms to recognize and interpret speech, and can support a wide range of voice-enabled experiences. AVS also includes a sentiment analysis feature that can identify the emotional tone of a user's voice. This feature can be used to personalize responses, adapt the device's behavior to the user's mood, and improve the overall user experience. AVS includes several pre-built models for sentiment analysis, which are trained on large datasets of speech samples with labeled emotional states [2, 3]. It also includes several features for managing user privacy, such as user opt-in, encryption, and secure communication protocols [4].

Amazon Comprehend : Amazon Comprehend is a natural language processing service provided by Amazon Web Services (AWS) that enables developers to analyze text and speech for sentiment and other language features. Comprehend uses deep learning models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to identify sentiment in text and speech data. It can detect positive, negative, and neutral sentiment, and can also identify the strength and intensity of the sentiment. Additionally, it can be trained on custom datasets to improve its accuracy for specific domains or applications. Comprehend can be used to analyze customer feedback, social media posts, and other forms of text data to gain insights into customer sentiment and opinions [5, 6].

Developers can use Comprehend to improve their products and services, track brand reputation, and identify emerging trends in customer feedback. Comprehend supports several languages, including English, Spanish, French, German, Italian, and Japanese[5, 6].

Google Nest uses the following technologies:

Google Cloud Natural API : The Google Cloud Natural Language API is a machine learning-powered service provided by Google Cloud. The Natural Language API uses deep learning models, including neural networks and convolutional neural networks, to analyze text and speech data for sentiment [7, 8, 9]. It has the capability to detect positive, negative, and neutral sentiment, and can also identify specific emotions, such as anger, joy, or sadness. It provides sentiment analysis capabilities for text data, along with other natural language processing features such as entity recognition and syntax analysis [7, 8, 9].
The Natural Language API can be customized with custom classifiers, which are trained on labeled datasets to improve accuracy for specific domains or applications.
The Natural Language API supports several languages, including English, Spanish, French, German, Italian, and Japanese [7, 8, 9].

Google DialogFlow: Google DialogFlow is a conversational AI platform provided by Google Cloud that enables developers to build chatbots and other conversational interfaces. It uses a combination of rule-based and machine learning-based models to interpret natural language input and generate appropriate responses. DialogFlow includes a sentiment analysis feature that uses machine learning models to detect the emotional tone of user input. This feature can be used to personalize responses, adapt the chatbot's behavior to the user's mood, and improve the overall user experience. It can be customized with custom classifiers, which are trained on labeled datasets to improve accuracy for specific domains or applications [10, 11].
DialogFlow supports several platforms and channels, including voice assistants, chatbots, messaging apps, and telephone systems [10, 11].

4. Connections in AI and Ethical Issues

To frame the understanding of sentiment analysis and its possible future extensions, an interview was conducted with a data scientist developing Augmented Reality and Virtual Reality (AR and VR). Presently, AR and VR are, relative to other platforms, in their infancy. Consequently, the present focus is on growing the user base. As it stands currently, AR/VR platforms are bug riddled and primarily used by niche gamers. Besides the obvious priority of developing a more functional, less bug-prone platform, sentiment analysis and other forms of user feedback are used to inform the structure that the AR/VR world should take. They want to track who spent how much time in

what part of the AR/VR platform – furthermore, did they enjoy their time, and did they come back? It is not yet time to deploy sentiment analysis in the service of product and promotional content recommendations. Per the data scientist, there is not yet enough good quality content to be able to use sentiment analysis — low sample size for users, reviews, and content.

This section of AI sounds innocuous, but the data scientist pointed out the parallel to Instagram. It was a small, $1B company when it was acquired by Facebook, that seemed destined to serve only a small subset of young people. However, it is now a colossus (> 2 Billion monthly active users) that precisely tracks user preferences as well as their comments and posts. Sentiment extraction from these interactions are used to determine the recommendations and promotional content that a user sees. Consequently, sentiment analysis and its associated deep learning methods are actively shaping the content recommendations for a third of the planet on a monthly basis on this single platform. Similarly, it is expected that the AR/VR will eventually, though it has not yet, make use of sentiment analysis.

Turning to a near term extrapolation, the data scientist describes the driving of new users to AR/VR platforms as the user interface improves and big content partnerships like the NBA and NBC incentivize adoption. Within this AR/VR realm, the spoken word, rather than text, is the operative medium of expression and communication. Consequently, the data scientist sees this as an even richer input source for the eventual deployment of sentiment analysis in service of Ad serving — tone, cadence, inflection and the words spoken can provide a higher resolution insight into consumer sentiment. This will have benefits for consumers who will be able to be served with more relevant ads and recommendations.

The data scientist sees some troubling questions concerning AR/VR as it gains market share:
1. How should behavior around the spoken word in a virtual world be policed? What are the appropriate internal and external strictures that should be used to define acceptable behavior?
2. How is the usage of the AR/VR platform being monitored for children and adolescents? Children, teens and young adults will represent most of AR/VR users – how is data surrounding both sensitive information and sentiment being collected and used?

A central problem, as the data scientist sees it, is the mismatch in speed between technological development and legislative action. This creates a gap whereby new technologies and platforms are left to make ethical decisions as they see fit.

There are varying approaches, on a national and international level, to consumer data privacy. Between China, the U.S. and the E.U., there are many varying approaches to consumer data privacy [14]. Interestingly, it is the E.U. (GDPR) and China (PIIPL, DSL) that have formally instantiated country/member-wide legislation that seeks to give privacy and data sovereignty back to the individual. Though China is still surveilling its citizens, it wants to minimize the ability of private – and especially foreign companies – to gather, transmit and analyze its citizens' data. Furthermore, China has even sought to regulate how predictive algorithms can be used to recommend content to Chinese consumers [15]. The U.S., on the other hand, has a patchwork of laws that are only applied in specific settings (e.g HIPAA, COPPA), but lacks any comprehensive laws to regulate most of the data that is currently being gathered, processed, and sold.

The steps taken by both China and the E.U. show two governments that are wildly different in their ideologies that have both nonetheless made similar conclusions; for data rights and privacy, the consumer is in need of comprehensive protections from a central authority. Taking these steps has, for China and the E.U., likely been easier when compared to the U.S. Their governments are not subject to the same political influence from large tech firms that the U.S. struggles with. Nonetheless, as aspirational regulators, the hope is to pass new legislation that more appropriately balances innovation with consumer data privacy.

Data privacy is not the only issue. Even as sentiment analysis becomes increasingly accurate and effective, biases still exist that cause problems. Sentiment analysis models are black boxes that take in data and output answers with few ways for humans to understand how they came to those conclusions. This can lead to biases in the model that are hard to find and harder to fix. A 2018 study showed that sentiment analysis of customer reviews was less effective for men than it was for women [11]. The article presumes that this is because men are less explicit in their reviews. This means that women's opinions on a product are given more weight, which could significantly bias a company's view of their products. This is only one of many biases that these models can have that has the potential to negatively affect a company's insights.

5. Regulation

Presently, outside of a few protected areas (healthcare, student records, etc.) the consumer in the U.S. has no ability to control how their data is processed, stored and sold. Though sentiment analysis is the specific area being regulated, it falls into a broader category of machine

learning deployments that seek to use data that has not necessarily been taken or used with full consent of the user. Accordingly, we are therefore speaking to the ethics and consumer welfare of any ML methods that use consumer data to, after feature extraction, produce actionable output — for instance who is feeling what, given a search query, and therefore more receptive to particular messages.

In order to successfully integrate sentiment analysis and other ML deployments, we must overcome two significant obstacles:
1. the pace at which new legislation is introduced and ratified
2. the influence that big tech firms represent through lobbying.

The second item is particularly difficult as, in the U.S, a prospective bill would need to pass with a majority in both the House and Senate. Consequently, there would be a battle to overcome the influence of large tech firms who would do their best to strangle any regulation that would reduce their ability to monetize consumer data. Nonetheless, the optimistic hope is that it is possible to overcome this issue and champion an overarching bill in the spirit of the E.U.'s General Data Protection Regulation. As the conversation with the AR/VR developer has shown, there is little present guidance when it comes to censorship, privacy and recommendations – it is mostly up to the companies implementing these platforms to devise their own codes of conduct. Therefore, it is imperative to devise a comprehensive, 21st century regulatory framework; one with the requisite flexibility to respond to technological developments. Outdated laws from the 80's continue to define the regulatory landscape, but it is important that any prospective regulation be broadly framed so that it may include the possible technical extensions going forward into the future.
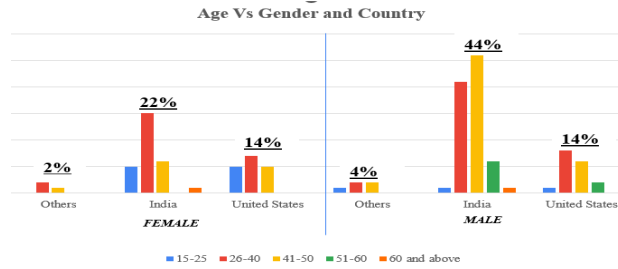
6. Conclusion:

As technology has advanced at an unprecedented pace in the past 30 years, laws and regulations have lagged behind. This inability to keep up with emerging technologies has led to a culture of self-regulation for large corporations, and has created no incentive for doing the right thing for the consumer if it is at the cost of company profit. Sentiment analysis has become a powerful tool for understanding customer's opinions on products and services, but as it evolves, so does the need for proper regulation. China and the E.U. have shown that it is possible to regulate these new technologies, but the U.S. government – due in part to the large lobbying budgets of these companies – has dragged its feet. Efforts should be focused on developing ethical frameworks, and standards to address these concerns, fostering trust and accountability in the field of sentiment analysis.

Socially, it is crucial to mitigate biases in sentiment analysis models and ensure fairness, avoiding the reinforcement of stereotypes or discriminatory practices. Protecting user privacy and complying with privacy regulations are essential considerations. From a managerial standpoint, sentiment analysis provides valuable insights, but responsible interpretation and consideration of ethical implications are necessary.By promoting responsible use and regulation, it is possible to maximize the potential of sentiment analysis while mitigating its pitfalls, ultimately leading to a more reliable, fair, and beneficial application of this technology.
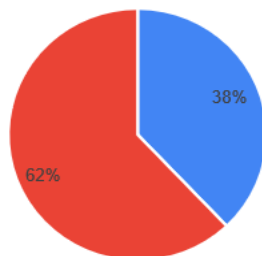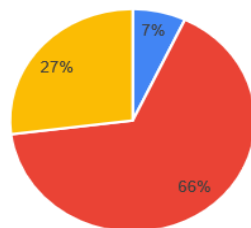
Appendixes

Appendix 1 – Demographics

Age Vs Gender and Country

15-25 26-40 41-50 51-60 60 and above
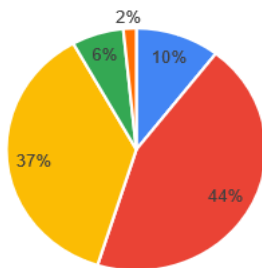
Age Distribution

15-25 26-40 41-50 51-60 60 and above

Gender Distribution

Female Male

Country Distribution

Other India United States
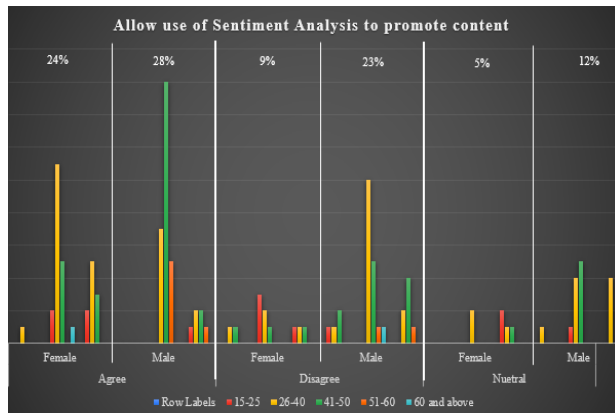
Appendix 2 – Embrace Technology Advancement Vs Willingness to Sacrifice Privacy

Improvements due to Technology advancement
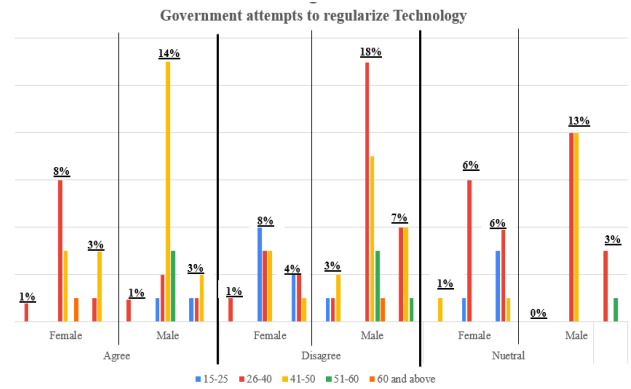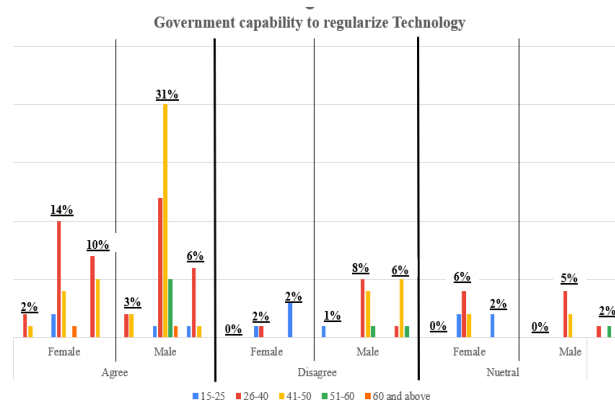

Agree Nuetral Disagree
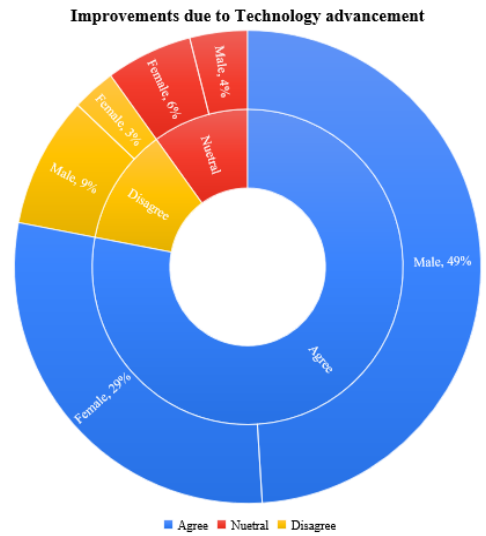
Willingness to sacrifice privacy

Agree Nuetral Disagree

Appendix 3 – Acceptance Levels of Sentiment Analysis

Sentiment Analysis does not waarrant any additional data when compared to speech and text

Allow use of Sentiment Analysis to promote content
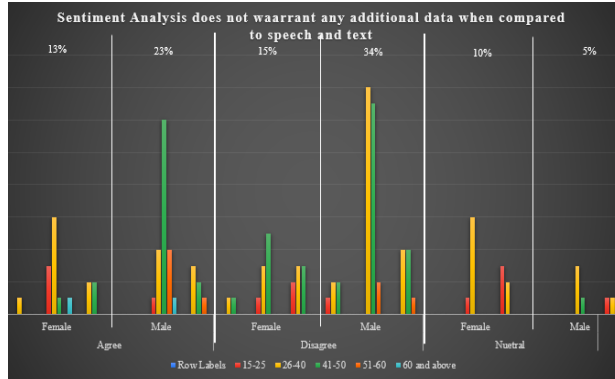
Appendix 4 – Respondents view on Government Role

Government capability to regularize Technology

Government attempts to regularize Technology

Appendix 5 – Product Trustworthiness
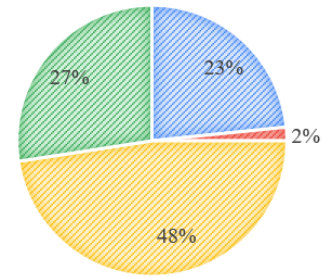
References

[1] Global Industry Analysts, 6361267 Global Sentiment Analytics Industry (reportlinker.com), March 2023.

[2] "Alexa Built-in Overview | Alexa Voice Service." Amazon, 28 Sept. 2022, developer.amazon.com/en-US/docs/alexa/alexa-voice-service/get-started-with-alexa-voice-service.html.

[3] "What Is the Alexa Skills Kit? | Alexa Skills Kit." Amazon (Alexa), developer.amazon.com/en-US/docs/alexa/ask-overviews/what-is-the-alexa-skills-kit.html.

[4] "Understand Smart Home Security Skills | Alexa Skills Kit." Amazon (Alexa), developer.amazon.com/en-US/docs/alexa/device-apis/overview-smart-home-security.html.

[5] "Amazon Comprehend - Natural Language Processing (NLP) and Machine Learning (ML)." Amazon Web Services, Inc., 2019, aws.amazon.com/comprehend/.

[6] "What Is Amazon Comprehend? - Amazon Comprehend." Docs.aws.amazon.com, docs.aws.amazon.com/comprehend/latest/dg/what-is.html.

[7] "Cloud Natural Language." Google Cloud, cloud.google.com/natural-language.

[8] "Cloud Natural Language Documentation | Cloud Natural Language API." Google Cloud, cloud.google.com/natural-language/docs.

[9]   "Dialogflow | Google Cloud." Google Cloud, 2019, cloud.google.com/dialogflow/.

[10] "Dialogflow Documentation." Google Cloud, cloud.google.com/dialogflow/docs.

[11] Thelwall, Mike. "Gender Bias in Sentiment Analysis." Emerald Insight, vol. 42, no. 1, 12 Feb. 2018, www.emerald.com/insight/content/doi/10.1108/OIR-05-2017-0139/full/html#sec007.

[12] "Sentiment Analysis." Bain, 8 Sept. 2020, www.bain.com/insights/customer-experience-tools-sentiment-analysis/. Accessed 1 May 2023.

[13] Bhavan, A., Sharma, M., Piplani, M., Chauhan, P., & Shah, R. R. (2020). Deep learning approaches for speech emotion recognition. In *Deep learning-based approaches for sentiment analysis* (pp. 259–289). Springer, Singapore

[14] Wolford, B. (2022, May 26). *What is GDPR, the EU's new Data Protection Law?* GDPR.eu. Retrieved May 2, 2023, from https://gdpr.eu/what-is-gdpr/

[15] Sponselee, Annika. Deloitte. (2022, August 25). *The Personal Information Protection Law (PIPL): China's equivalent to the GDPR?* Deloitte Netherlands. Retrieved May 2, 2023, from https://www2.deloitte.com/nl/nl/pages/risk/articles/pipl-chinas-equivalent-to-the-gdpr.html